

TW 333A

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
21 April 2005 (21.04.2005)

PCT

(10) International Publication Number
WO 2005/036405 A1

(51) International Patent Classification⁷: G06F 12/10

(21) International Application Number:
PCT/US2004/033253

(22) International Filing Date: 7 October 2004 (07.10.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/509,581 8 October 2003 (08.10.2003) US

(71) Applicant (for all designated States except US): UNISYS CORPORATION [US/US]; Unisys Way, MS/E8-114, Blue Bell, PA 19424-0001 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LANDIS, John, A. [US/US]; 7124 Old Easton Road, Pipersville, PA 18947 (US). POWDERLY, Terrence, V. [US/US]; 10 Crown Lane, East Fallowfield, PA 19320 (US). SUBRAHMANIAN, Rajagopalan [IN/US]; 731 Parkview Drive,

Phoenixville, PA 19460 (US). PUTHIYAPARAMBIL, Aravindh [IN/US]; 715 Parkview Drive, Phoenixville, PA 19460 (US). HUNTER, James, R., Jr. [US/US]; 808 Bush Lane, Chadds Ford, PA 19317 (US).

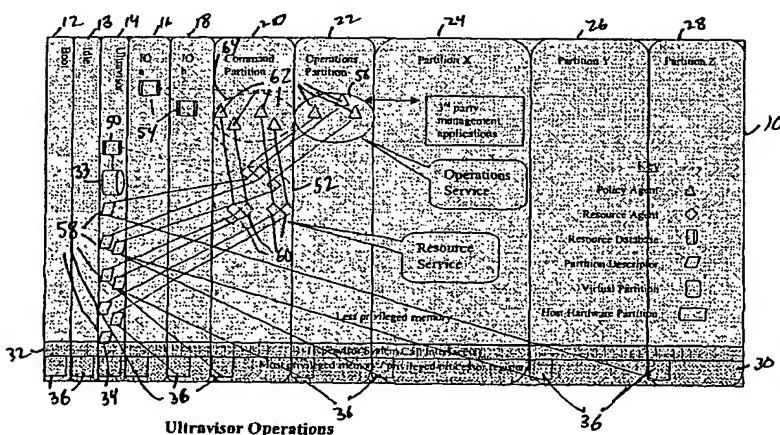
(74) Agents: STARR, Mark, T. et al.; Unisys Corporation, Unisys Way, MS/E8-114, Blue Bell, PA 19424-0001 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),

[Continued on next page]

(54) Title: COMPUTER SYSTEM PARA-VIRTUALIZATION USING A HYPERVISOR THAT IS IMPLEMENTED IN A PARTITION OF THE HOST SYSTEM



(57) Abstract: A virtualization infrastructure that allows multiple guest partitions (24, 26, 28) to run within a host hardware partition (10). The host system is divided into distinct logical or virtual partitions (24, 26, 28) and special infrastructure partitions (12-14, 16, 18, 20, 22) are implemented to control resource management and to control physical I/O device drivers that are, in turn, used by operating systems in other distinct logical or virtual guest partitions (24, 26, 28). Host hardware resource management runs as a tracking application in a resource management "ultravisor" partition (14), while host resource management decisions are performed in a higher level command partition (20) based on policies maintained

in a separate operations partition (22). The conventional hypervisor (32) is reduced to a context switching and containment element (monitor) for the respective partitions (24, 26, 28), while the system resource management functionality is implemented in the ultravisor partition (14). The ultravisor partition (14) maintains the master in-memory database (33) of the hardware resource allocations and serves a command channel to accept transactional requests for assignment of resources to partitions (24, 26, 28). It also provides individual read-only views of individual partitions (24, 26, 28) to the associated partition monitors. Host hardware I/O management is implemented in special redundant I/O partitions (16, 18). Operating systems in other logical or virtual partitions (24, 26, 28) communicate with the I/O partitions (16, 18) via memory channels established by the ultravisor partition (14). The guest operating systems in the respective logical or virtual partitions (24, 26, 28) are modified to access monitors that implement a system call interface through which the ultravisor (14), I/O (16, 18), and any other special infrastructure partitions (12-13, 20, 22) may initiate communications with each other and with the respective guest partitions (24, 26, 28). The guest operating systems are modified so that they do not attempt to use the "broken" instructions in the x86 system that complete virtualization systems must resolve by inserting traps.

BEST AVAILABLE COPY

WO 2005/036405 A1



European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.